Computer Methods and Programs in Biomedicine 157 (2018) 163–177

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

dfpk: An R-package for Bayesian dose-finding designs using pharmacokinetics (PK) for phase I clinical trials

A. Toumazi[a], E. Comets[b,c], C. Alberti[d], T. Friede[e], F. Lentz[f], N. Stallard[g], S. Zohar[a,1], M. Ursino[a,1,*]

[a] INSERM, UMRS 1138, Team 22, CRC, University Paris 5, University Paris 6, Paris, France
[b] INSERM, CIC 1414, University Rennes-1, Rennes, France
[c] INSERM, IAME UMR 1137, University Paris Diderot, Paris, France
[d] INSERM, UMR 1123, Hôpital Robert-Debré, APHP, University Paris 7, Paris, France
[e] Department of Medical Statistics, University Medical Center Göttingen, Göttingen, Germany
[f] Federal Institute for Drugs and Medical Devices, Bonn, Germany
[g] Statistics and Epidemiology, Division of Health Sciences, Warwick Medical School, The University of Warwick, UK

ARTICLE INFO

Article history:
Received 30 June 2017
Revised 11 January 2018
Accepted 24 January 2018

Keywords:
Dose-finding
Maximum tolerated dose
Pharmacokinetics
Phase I clinical trials
R package

ABSTRACT

Background and objective: Dose-finding, aiming at finding the maximum tolerated dose, and pharmacokinetics studies are the first in human studies in the development process of a new pharmacological treatment. In the literature, to date only few attempts have been made to combine pharmacokinetics and dose-finding and to our knowledge no software implementation is generally available. In previous papers, we proposed several Bayesian adaptive pharmacokinetics-based dose-finding designs in small populations. The objective of this work is to implement these dose-finding methods in an R package, called `dfpk`.

Methods: All methods were developed in a sequential Bayesian setting and Bayesian parameter estimation is carried out using the `rstan` package. All available pharmacokinetics and toxicity data are used to suggest the dose of the next cohort with a constraint regarding the probability of toxicity. Stopping rules are also considered for each method. The `ggplot2` package is used to create summary plots of toxicities or concentration curves.

Results: For all implemented methods, `dfpk` provides a function (`nextDose`) to estimate the probability of efficacy and to suggest the dose to give to the next cohort, and a function to run trial simulations to design a trial (`nsim`). The `sim.data` function generates at each dose the toxicity value related to a pharmacokinetic measure of exposure, the AUC, with an underlying pharmacokinetic one compartmental model with linear absorption. It is included as an example since similar data-frames can be generated directly by the user and passed to `nsim`.

Conclusion: The developed user-friendly R package `dfpk`, available on the CRAN repository, supports the design of innovative dose-finding studies using PK information.

1. Introduction

Dose-finding studies and pharmacokinetics (PK) are carried out at the first phases of clinical evaluation of a new drug in humans. Drug safety is evaluated in the dose-finding study, which aims at identifying the maximum tolerated dose (MTD) [1]. Meanwhile, the PK data collected during such study provides the description of the dose-concentration relationships [2]. Nevertheless, these two approaches are often conducted and reported independently in different sections in publications reporting trial results [3]. Identifying the right dose or set of doses at an early stage is crucial: selecting too toxic doses can result in patient overdosing, while selecting an inefficient dose increases the likelihood that the drug will be found to be ineffective in subsequent clinical evaluation [4]. Particularly in the case of small populations, such as rare diseases or paediatrics, it should be useful to take into account all the infor-

* Corresponding author..
E-mail address: moreno.ursino@inserm.fr (M. Ursino).
[1] The last two authors contributed equally as co-senior authors.

https://doi.org/10.1016/j.cmpb.2018.01.023

mation collected during the trial, and to try to utilize the PK measurements within the dose-finding design. Only few attempts have been described in the literature so far and, usually, the methods were built for a very specific situation [5–8]. Moreover, no software implementations are publicly available.

In this article we present the new R package `dfpk` (short for dose-finding Pharmacokinetics), which provides the Bayesian adaptive PK-based dose-finding designs in small populations proposed by Ursino et al. [8] through the freely available R software [9]. The six methods detailed in [8] have been implemented in `dfpk`. For each of them, two functions are provided: (i) a function to determine the next recommended dose (during the trial) or the recommended MTD (at the end of the trial) and (ii) a function to run simulations of phase I studies to design a new trial. Interactive graphical representations of the dose-concentration curve, of the dose allocation process in the trial and of the dose-toxicity response are also provided by the package.

The paper is organised as follows. Section 2 introduces the statistical methods proposed by Ursino et al. [8], along with the description of the suggested scenarios to be simulated. Section 3 outlines the structure of the package and the main functions of the paper (`sim.data`, `nextDose` and `nsim`) with practical examples. Section 4 & 5 include conclusion, discussion and recommendations.

2. Computational methods

The present section briefly reviews the methods proposed by Ursino et al. [8] to perform dose-finding taking into account PK measurements. We then describe the scenarios simulated in [8] in order to evaluate the robustness of the method, which have been added as examples in the dfpk package.

2.1. Dose-finding methods

Let $D = \{d_1, \ldots, d_k\}$ be the set of K possible doses with $d_1 < \cdots < d_k$ and $d_{[i]} \in D$ be the dose administered to the ith subject ($i = 1, \ldots, n$, where n denotes the sample size) and y_i be a binary variable which takes value 1 if the ith subject shows a DLT (dose-limiting toxicity) and 0 otherwise. Moreover, let z_i be the logarithm of the area under the curve (AUC) of the concentrations of drug in blood plasma against time, for the ith patient.

All methods share the same fundamental idea for the dose-escalation rule: the dose chosen for the next cohort enrolled is the one with probability of toxicity nearest to the target θ selected by the trial investigators. A no-skipping rule is given: if some doses have not yet been tested, the dose is chosen from $D^* \subset D$, a subset of D which contains all the doses already evaluated and the first dose level immediately above. The final recommended MTD is given by the dose that would have been administered for the $(n+1)$st subject enrolled in the trial. Finally, we added in all methods the same stopping rule: if the posterior probability of toxicity of the first dose is greater of a specified threshold, then no dose is suggested and the trial is stopped.

Each method is separated from the others. We adopted the convention of starting the subscription of β parameter from 0 for each method. Therefore, even if the parameters share the same names across models, they have different interpretations. In the following, we briefly describe how the probability of toxicity is estimated and computed in each method.

2.1.1. PKCOV

PKCOV is a modification of the method proposed by Piantadosi and Liu [5] who suggested to use the AUC as a covariate for p_T, the probability of toxicity, through the logit link. Therefore, the dose-toxicity model is

$$\text{logit}(p_T(d_k, \Delta z_{dk}, \beta)) = -\beta_0 + \beta_1 \log(d_k) + \beta_2 \Delta z_{dk} \quad \forall d_k \in D, \tag{1}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2)$, β_0 is a constant selected through a sensitivity analysis or with prior information, Δz_{dk} is the difference between the logarithm of population AUC at dose d_k and z, the logarithm of AUC of the subject at the same dose. Independent uniform distributions are selected as prior distributions for β_1 and β_2. In detail, $\beta_1 \sim U(\max(0, m_1 - 5), m_1 + 5)$, where m_1 reflects the prior information on the parameter and the length of the domain of the distribution can go up to 10, and $\beta_2 \sim U(0, 5)$. Both β_0 and m_1 should be selected using prior information, such as from preclinical data, and sensitivity analysis should be done. The estimated probability of toxicity versus dose is obtained by inverting Eq. (1), using $\beta_1 = \hat{\beta}_1$, the estimated parameter, and $\Delta z_{dk} = 0$.

2.1.2. PKLIM and PKCRM

PKLIM is a modification of the method proposed by Patterson et al. [6] and Whitehead et al. [10]. A normal PK-toxicity model is used:

$$z_i | \boldsymbol{\beta}, \nu \sim N\left(\beta_0 + \beta_1 \log d_i, \nu^2\right), \tag{2}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1)$ are the regression parameters, and ν is the standard deviation. A bivariate normal distribution and a beta distribution are chosen for $\boldsymbol{\beta}$ and ν, respectively, that is, $\boldsymbol{\beta} \sim N(\mathbf{m}, \nu^2(g\mathbb{1}))$ and $\nu \sim Beta(1, 1)$. Therefore, a hierarchical prior distribution is given to $\boldsymbol{\beta}$, where $\mathbf{m} = (-\log Cl_{pop}, 1)$ and g should be chosen using prior information. For instance, Cl_{pop} denotes the attended value of the clearance at population level, and g reflects the precision. The probability of toxicity of each dose is computed as

$$P(z > L \mid d_k, \boldsymbol{\beta} = \hat{\boldsymbol{\beta}}, \nu = \hat{\nu}) \quad \forall d_k \in D, \tag{3}$$

where L is a threshold set before starting the trial and the hat denotes the posterior means of the parameters. Since an assumption underlying the model is that DLTs are based on the AUC exceeding some threshold, the method could be applicable only when such a threshold is known. In order to avoid this problem, the PKCRM method was proposed, which is the combination of PKLIM and the CRM [11] using a power working model and normal prior on the parameter. In PKCRM the dose recommended for the next subject is the lowest of the doses recommended by the two methods.

Note that although the same notation has been used for convenience, the parameters β_0 and β_1 are different in the different models.

2.1.3. PKLOGIT, PKPOP, PKTOX

PKLOGIT, inspired by Whitehead et al. [7], combines two regressions to compute the probability of toxicity versus the dose. The first one is the same as Eq. (2), that is z versus dose. In the second, z is used as a covariate in a logistic regression model for p_T. This means that now the probability of toxicity is described in term of AUC and not any more in term of dose. Therefore, we have that

$$\text{logit}(p_T(z, \boldsymbol{\beta})) = -\beta_2 + \beta_3 z, \tag{4}$$

where β_2 and β_3 have independent uniform prior distributions, that is, $\beta_2 \sim U(0, m_2)$ and $\beta_3 \sim U(0, m_3)$, with $m_2 \geq m_3$, and values can be chosen using prior information. If no information is available, $m_2 = 20$ and $m_3 = 10$ are good starting values for a sensitivity analysis. The probability of toxicity associated with each dose is obtained by using the estimated parameters of each regression model in the following expected value formula:

$$P(y = 1 \mid d_k, \boldsymbol{\beta} = \hat{\boldsymbol{\beta}}, \nu = \hat{\nu}) = E\left[\frac{1}{1 + e^{\hat{\beta}_2 - \hat{\beta}_3 z}}\right] = \int \frac{1}{1 + e^{\hat{\beta}_2 - \hat{\beta}_3 z}} g(z)\, dz, \tag{5}$$